FINAL TECHNICAL REPORT

Earthquake Monitoring in an Urban Setting Through Deep-Learning-Based Noise Suppression
and Back-Projection: Case Study for the Long Beach Dense Array

Gregory C. Beroza and Lei Yang

Stanford University
Department of Geophysics
397 Panama Mall
Stanford, CA, 94305-2215
Telephone: (650)723-4746; fax: (650) 725-7344

Principal Investigator: Gregory C. Beroza (beroza@stanford.edu)

Abstract

Earthquake monitoring in urban settings is crucial but challenging due to the strong cultural noise in seismic recordings inherent to urban environments. Here we develop a deep-learning-based seismic denoising algorithm – UrbanDenoiser – to suppress the seismological urban noise. The algorithm is trained using a waveform data set that combines rich noise sources from the urban Long Beach dense array and high signal-to-noise ratio (SNR) earthquake signals from the rural San Jacinto dense array. Application of UrbanDenoiser to urban seismic data demonstrates that UrbanDenoiser can strongly suppress the seismic noise levels relative to the signals, and that the seismic signals can be recovered from noisy data with SNR floor around one. Earthquake location based on the denoised continuous Long Beach data does not support the previously reported observational result of mantle seismicity beneath Los Angeles, while it suggests a fault model featuring shallow creep, intermediate locking, and localized stress concentration at the base of the seismogenic zone.

Introduction

Earthquake risk is highest in urban settings due to both population density and to the presence of extensive and vulnerable infrastructure. Ideally, intensive earthquake monitoring efforts in urban areas would be used to characterize the fault systems that pose the most immediate and direct threats to cities; however, the same factors – population and infrastructure – that cause the risk exposure to be high, also make earthquake monitoring difficult to carry out due to both the various kinds of seismic noise generated in cities, and to the logistical difficulties of instrumental deployments.

The Los Angeles metropolitan area is located within an active plate boundary. The Newport-Inglewood Fault runs directly through Los Angeles (Fig. 1), as do other faults that either traverse it, or bound the Los Angeles Basin, including the Palos Verdes, Santa Monica-Hollywood, Sierra Madre, Whittier Faults and some blind faults. Microseismic monitoring is important for this densely populated city because earthquake locations provide essential constraints on the location and geometry of active faults and the hazards they pose[1-3].
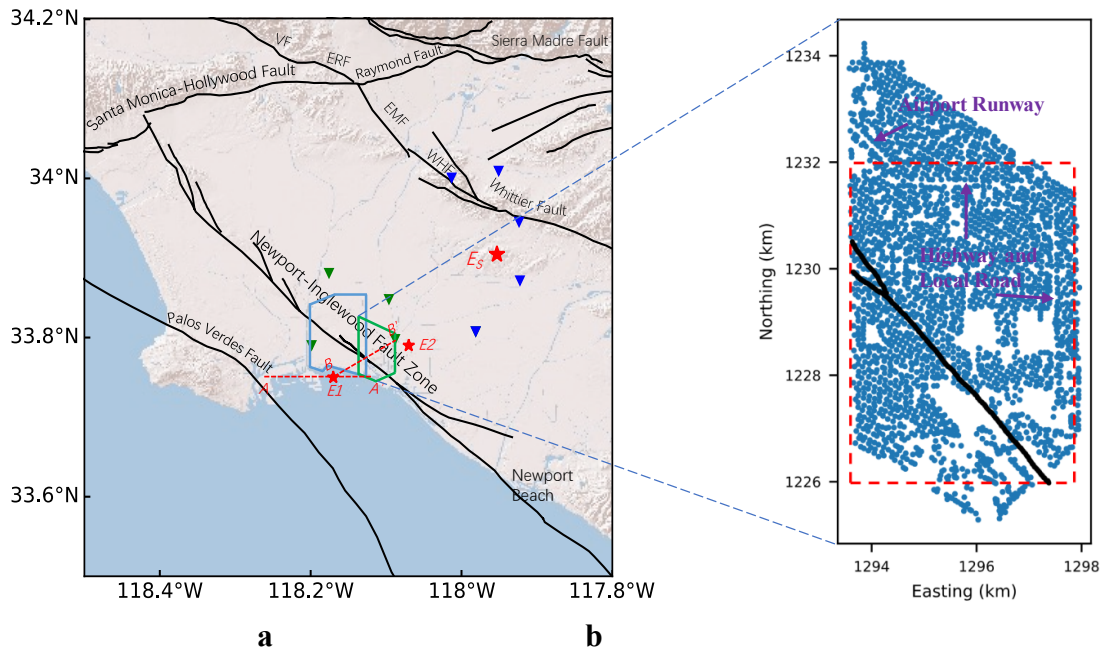

Figure 1. Los Angeles Basin and Long Beach dense nodal deployment. a Map of the Los Angeles Basin showing Newport-Inglewood and other faults (VF: Verdugo Fault; ERF: Eagle Rock Fault; EMF: East Montebello Fault; WHF: Workman Hill Fault). Blue and green polygons outline the Long Beach phase A and B, deployments. Red stars *E1* and *E2* show epicenters of earthquakes that occurred during the deployment of Long Beach phase B. *AA'* is a profile across the epicenter *E1*. *BB'* is a line across the epicenter *E1*. Red star E_S shows the main shock epicenter of 2014 La Habra earthquake sequence. Green and blue triangles show regional stations from SCSN close to the Long Beach deployment and the earthquake sequence, respectively. **b** Map of Long Beach phase B deployment. Stripe of missing sensors in the upper left is the Long Beach Airport runway. Narrow gap in the northern and eastern part of the deployment tracks the highway and local roads. Black lines show surface trace of the Newport-Inglewood Fault. Red dashed rectangle is the surficial boundary of the 3D imaging volume we use in our analysis.

Traditional earthquake monitoring methods that use single-station measurements to detect wave arrivals for events above the noise floor on individual channels may fail to detect smaller events[4-6]. Dense array data provide an opportunity to detect and analyze these weak sources because adjacent stations have common signal attributes that can be exploited for detection. In 2011 and 2012, dense arrays with ~100 m spacing were deployed in two phases in Long Beach (Fig. 1a, blue and green polygons). Phase A (blue polygon) operated for the first six months of 2011, covered an area of 10 km × 7 km, and included approximately 5200 vertical 10 Hz geophones with sampling frequency of 500 Hz. Phase B (green polygon) extended the original survey towards the east and shared a similar tectonic setting above a branch of the Newport-Inglewood Fault Zone (NIFZ). It operated for the first three months of 2012, covered an area of 8.5 km × 4.5 km, and included approximately 2,500 geophones.

Inbal et al.[7, 8], Li et al.[9] and Yang et al.[10] used the density of seismic wavefield data from these Long Beach arrays for microseismic monitoring. To suppress the strong cultural noise from the Long Beach Phase A data, Inbal et al.[7,8] used downward continuation to back propagate the wavefield recorded at surface to 5 km depth and performed back-projection (BP) to locate/image seismic events below that. They detected and located widespread seismicity at depths greater than 20 km in the upper mantle, which is much deeper than the conventionally determined and widely accepted seismogenic depth limit of continental earthquakes in this region[11]. Li et al.[9] used local waveform similarity to detect small events from the low signal-to-noise ratio (SNR) data. Their finding differs from Inbal et al.'s in that they only detected events with shallow origins. Yang et al.[10] applied a trace-randomization procedure to assess the reliability of the upper mantle earthquakes using back-projected Long Beach Phase B data. By comparing the seismic location results between the original and trace-randomized data, they inferred that the deep upper-mantle events found by Inbal et al may not be reliable event detections.

The discrepancy among these results occurs primarily due to the low SNR of the data. While methods like downward continuation/BP and local waveform similarity can decrease the detection threshold for small earthquakes, they are sensitive to both noise and uncertainties in the velocity structure. Seismic denoising has the potential to enhance detection sensitivity and has the flexibility to improve results for a broad range of approaches to earthquake detection/location as well as for seismic structural imaging[12-14].

Traditional denoising methods based on simple spectral filtering fail when seismic signals and noise overlap within the same frequency band. Time-frequency domain denoising can overcome this problem, however, the choice of a suitable thresholding function to map the noisy data into optimally denoised signal is challenging. Machine learning techniques, especially deep learning[15] provide a powerful approach to learn complex functional relationships and to use them to extract useful characteristics from very large data sets[16-19]. It provides a promising approach for time-frequency denoising methods through sparse representation of data and improved signal versus noise. Zhu et al.[20] developed DeepDenoiser based on a deep neural network, which significantly improves the SNR with minimal changes in the waveform shape of interest. DeepDenoiser was originally trained on an extensive dataset from Northern California that were recorded on instruments deployed in unpopulated, low-noise settings. DeepDenoiser effectively denoises independent seismic data recorded in that setting but did not generalize well to the Long Beach dataset, presumably because the noise sources differ from those of Northern California seismic dataset that the network was trained on.

The Long Beach dataset represents a rich data source of seismological urban noise[21]. In this paper, we develop a machine learning-based denoising method by exploiting this rich noise

resource within the framework of DeepDenoiser to suppress the strong noise level for seismic data recorded in the urban setting. To explore the validity of the previously reported widespread seismicity down to the upper mantle beneath Long Beach, we specifically include high SNR seismic signals from the San Jacinto dataset[5] in the training dataset for the neural network to learn the seismic signature of real earthquakes recorded on the same instruments, but in a quieter setting. We demonstrate that this deep-learning-based denoising tool has the potential to improve the detection capability of earthquake monitoring networks in urban settings.

Results

Network training

We develop UrbanDenoiser by training the deep neural network with seismic noise from the Long Beach dataset and seismic signals from the San Jacinto dataset. The architecture of the neural network is based on that of the DeepDenoiser algorithm[20]. The dataset consists of 90-s windows of seismic waveforms for 80,000 noise samples and 33,751 signal samples. The signal and noise samples are randomly split into training and validation sets. We generate noisy waveforms at different SNR levels by combining the signal training set repeatedly with randomly selected noise samples from the noise training set, and randomly shifting the waveform in the window[22]. The input for the neural network is the 2D time-frequency representation of noisy waveforms determined by Short Time Fourier Transform. Both the real and imaginary parts are input into the neural network so that it is able to learn from the time and phase information. The prediction targets are two masks for recovered signal and noise respectively. We generate seismic waveforms for the validation set with the same procedure and apply them for fine-tuning the hyper-parameters of the network. We test the neural network with the additional seismic data from Long Beach seismic recordings.

We extract noise samples from Long Beach data. These waveforms include various kinds of traffic sources (cars, airplanes, helicopters), vibroseis events and other unknown activities[23]. We collect seismic recordings from all the receivers in the Long Beach Phase B deployment on Julian days 27 and 48, 2012, and select seismic noise samples from them, because there are fewer earthquakes during these two days in the Quake Template Matching (QTM) catalog[24]. We segment the data in 90-s-long time series and remove those containing earthquake signals either from known seismic events in the QTM catalog or as determined by the PhaseNet algorithm[25].

The signal samples are extracted from San Jacinto dataset, which were recorded by another dense array deployed on the active Clark branch of San Jacinto Fault from 2014 May 7 to 2014 June 13[5]. This deployment consists of ~ 1,108 geophones that collected high-quality seismic signals from small to medium magnitude local earthquakes. The two deployments used the same sensors with the same instrumental response. We select the labeled signals with a strict condition. We run PhaseNet on the continuous data, and the candidate earthquake signal waveforms are selected based on their coherence across the seismic network. We select only those signal windows with SNR > 12 (defined as the root-mean-square ratio of the seismic energy after and before the first arrival) as the labeled signals. We also include 30,000 seismic signal samples from the North California Seismic Network in the training dataset to increase the predictive power of deep neural network and reduce overfitting.

Back-projection imaging for a local earthquake with denoised seismic data

A local M 2.1 earthquake on March 27, 2012 (red star E1 in Fig. 1a) was recorded by the Long Beach dense array. Figure 2a and b shows the corresponding seismic profiles for the original data and the data as denoised by UrbanDenoiser. The traces are sorted by source-receiver distance. Seismic signals in Fig. 2a are affected by strong noise, while after denoising, most of the seismic noise is suppressed in Fig. 2b.

We perform BP to image this earthquake (see methods below). Figure 2c-e show the BP imaging results along a 14-km W-E profile (red dashed line AA' in Fig. 1a) crossing the hypocenter with original data and denoised data by DeepDenoiser and UrbanDenoiser, respectively. Comparing Fig. 2c, d and e, we find that although the locations determined based on data processed by different procedures are similar, the peak amplitudes are more strongly affected, with the UrbanDenoiser result showing the largest amplitude, and the DeepDenoiser result showing the smallest amplitude. The imaging result based on UrbanDenoiser has the largest peak amplitude because the SNR of the seismic data is improved after denoising. The imaging result based on DeepDenoiser has the smallest peak amplitude, because the recovered signals lose part of the earthquake signal information.

We select 59 traces along a line crossing the epicenter and perpendicular to the surface trace of the Newport-Inglewood Fault (red dashed line BB' in Fig. 1a) and plot the denoised seismic waveforms in a profile sorted by source-receiver distance (Fig. 2f). The intersection of the fault trace and the line is closest to Trace 20. From Fig. 2f, we clearly see that the lineup for the first 20 traces has a faster move-out than for Traces 21-59. This indicates that the P-wave velocity to the west of the fault trace is faster than that to the east[26]. Note that the S-wave arrival can be clearly seen in the denoised data.
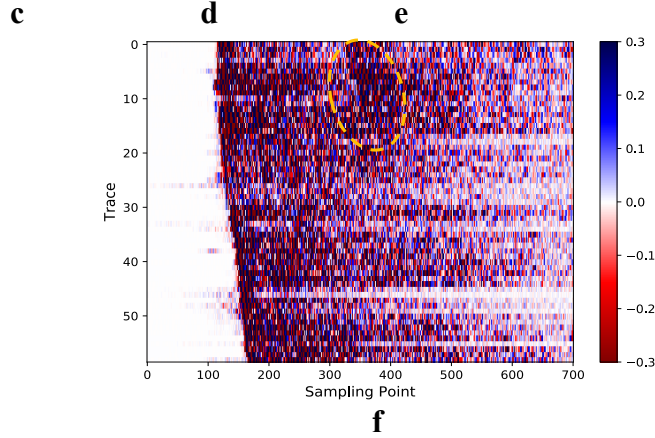

a   b
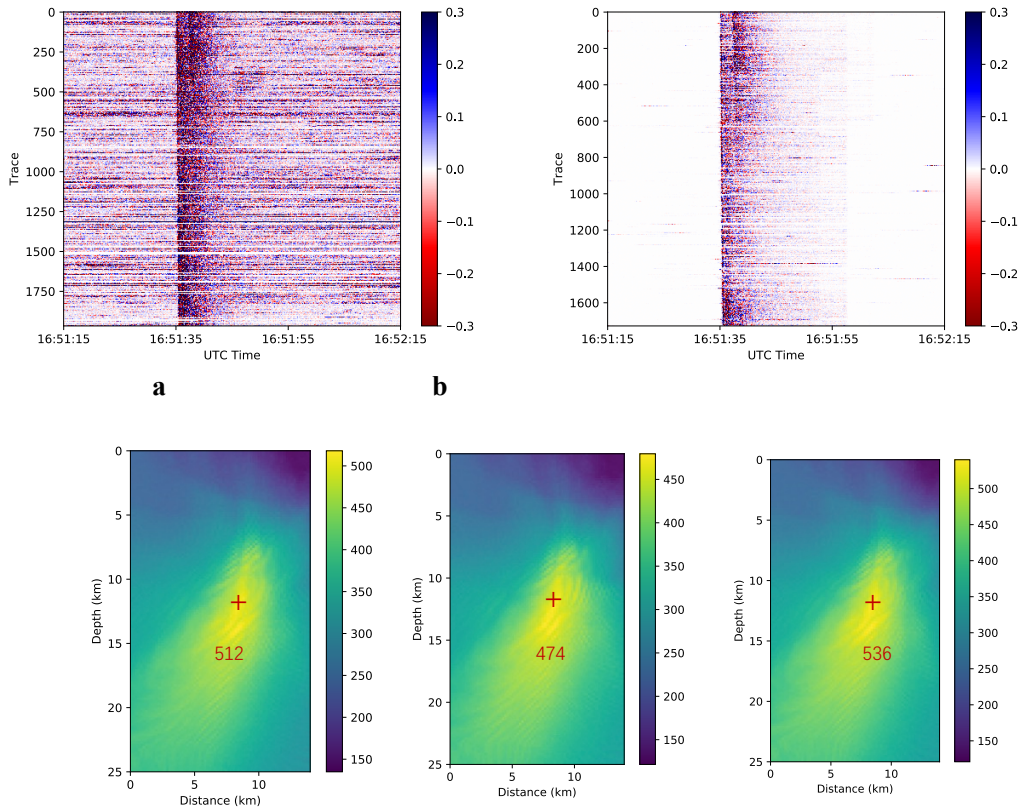
Figure 2. Seismic profile and back-projection imaging results for a local *M* 2.1 earthquake.
a Seismic profile on raw data. **b** Seismic profile on denoised data. The amplitude drop around 16:51:57 occurs because the neural network decision boundary between earthquake signal and noise implemented in the time-frequency domain. For the purpose of earthquake detection and location, the coda wave does not influence the result. **c** Back-projection imaging on raw data with the largest normalized amplitude of 512. **d** Back-projection imaging on denoised data by DeepDenoiser with the largest normalized amplitude of 474. **e** Back-projection imaging on denoised data by UrbanDenoiser with the largest normalized amplitude of 536. **f** Seismic profile containing 59 traces along a line crossing the epicenter and perpendicular to the surface trace of the Newport-Inglewood Fault; Orange ellipse marks the S wave arrival.

Earthquake Location with Denoised Long Beach Dense Array Data

We apply UrbanDenoiser to seven-days of seismic data (Julian days 61- 67) and perform BP on the denoised continuous data within a 4.4 ✗ 6.0 ✗ 25.0 km³ imaging volume, the boundary of which is shown as a red dashed rectangle in Fig. 1b, to detect and locate the most likely seismic sources. Robust seismic denoising allows us to work on the entire day's data, not just during the night when cultural noise is lower, as had been done previously[7, 10]. Figure 3a shows a one-day seismogram recorded by a randomly selected receiver that reveals a strong time-varying behavior in amplitude for local time 6:00 am – 10:00 pm (UTC time 14:00 – 6:00[+1]) when noise level is high vs. local time 10:00 pm – 6:00 am (UTC time 6:00 – 14:00) when noise level is low. Figure 3c and d show the denoised results from Fig. 3a and b, which eliminates the daytime/nighttime variation.
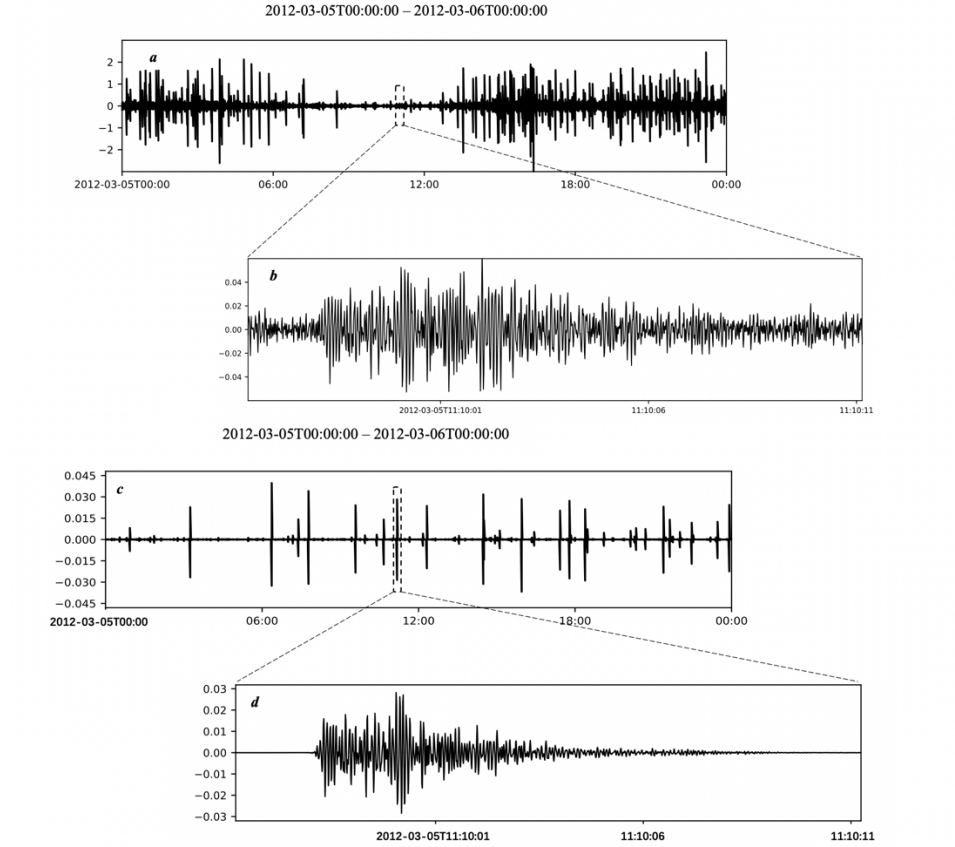
Figure 3. Seismogram recorded by station R1134_5043 and the denoised version (in UTC time). a One-day raw data. **b** Zoomed view of a microearthquake event in **a**. **c** Denoised version for **a**. **d** Denoised version for **b**. Some of the spikes in **c** are false positives, and we eliminate their influence in earthquake detection by utilizing the waveform coherence across the dense array.
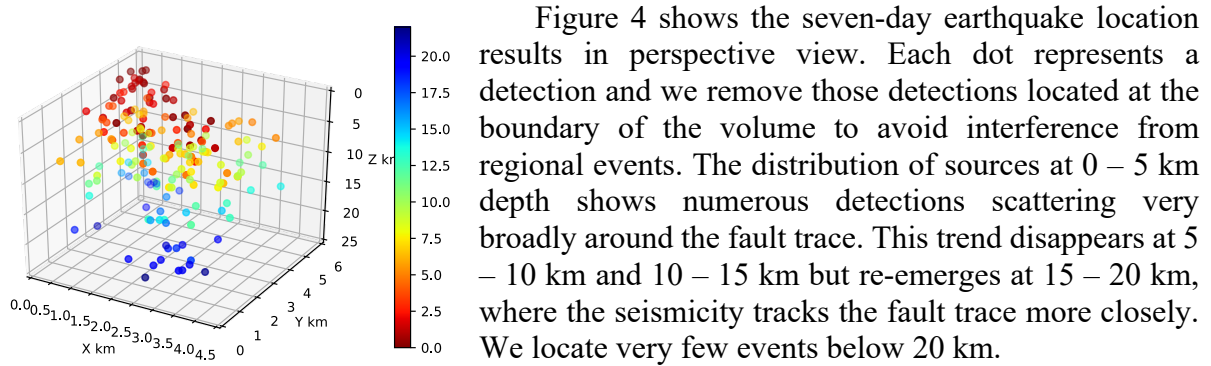

Figure 4 shows the seven-day earthquake location results in perspective view. Each dot represents a detection and we remove those detections located at the boundary of the volume to avoid interference from regional events. The distribution of sources at $0 - 5$ km depth shows numerous detections scattering very broadly around the fault trace. This trend disappears at $5 - 10$ km and $10 - 15$ km but re-emerges at $15 - 20$ km, where the seismicity tracks the fault trace more closely. We locate very few events below 20 km.

Figure 4. Seven-day earthquake BP location results.

We check our detection/location results by examining the seismic waveforms from the dense array dataset. We select one detection at (3.4 km, 0.8 km, 5 km) in Fig. 4a, and plot the seismic profiles spanning the duration of the earthquake. Figure 5a shows the seismic profiles on the raw Long Beach data, from which we can barely identify the seismic signals due to the strong background noise relative to the weak earthquake energy; however, after denoising, we can clearly

see the seismic arrivals in Fig. 5b. Figure 5c shows the zoomed view of the subset from b. The increasing first arrival time on traces sorted by the distances between each station and the determined epicentral location supports the validity of the location result. This detection is also validated by checking the seismograms recorded by the isolated regional stations from Southern California Seismic Network (SCSN).
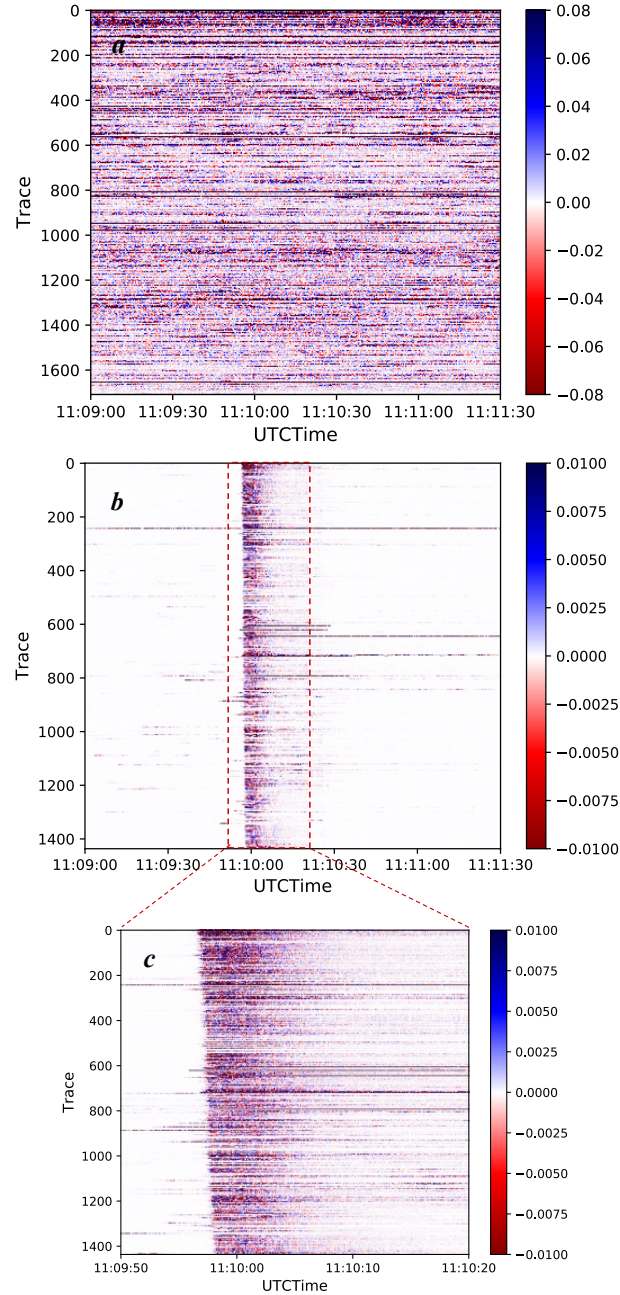

Figure 5. Seismic profile containing information for a small earthquake. a Raw data. **b** Denoised data. **c** Zoomed view for the red rectangle in **b**.

An earthquake on Julian day 67 is detected by the Long Beach nodal array. It occurred 2.5 km to the east of the imaging volume (Fig. 1a, red star *E2*), so the seismic energy is back-projected to a point on the east boundary of the imaging volume, which is excluded from Fig. 4.

Application of UrbanDenoiser to Regional Stations for an Earthquake Sequence

An earthquake sequence struck urban La Habra with a mainshock magnitude of 5.1 at 4:09:41 UTC on Mar. 29, 2014. We choose the five stations from SCSN nearest to the sequence (Fig. 1a, blue inverted triangles), and apply UrbanDenoiser to the seismograms. We confirm an earthquake when the detected phases can be associated on two or more stations, and by doing this we find a total of 488 events during the 10 hours between 3:00 ~ 12:00. This amount is 10% higher than the number in the QTM catalog.

Figure 6 shows 40-minute seismograms (03:20 – 04:00 UTC, Mar. 29, 2014, vertical component only) from the five stations. This is a period between the M 3.57 foreshock and M 5.1 main shock, and is a relatively quieter window compared with those following the main shock. The only event in the QTM catalog during this time is a M 0.67 earthquake at 3:40:59, which is also detected in the denoised waveforms shown in $a_{4\text{-}II}$-$d_{4\text{-}II}$. Comparing Fig. $6a_{4\text{-}II}$-$d_{4\text{-}II}$ with the raw data in Fig. $6a_{3\text{-}II}$-$d_{3\text{-}II}$, we find substantial enhancement of the SNR in the denoised version. With the denoised data, we find a total of nine events during this 40-minute period. Figures $6a_{4\text{-}I}$-$d_{4\text{-}I}$ and $6a_{4\text{-}III}$-$d_{4\text{-}III}$ show two examples (not included in QTM catalog) compared with the raw data in Figs. $6a_{3\text{-}I}$-$d_{3\text{-}I}$ and $6a_{3\text{-}III}$-$d_{3\text{-}III}$. This demonstrates that UrbanDenoiser can facilitate the detection of more small events in an urban setting.
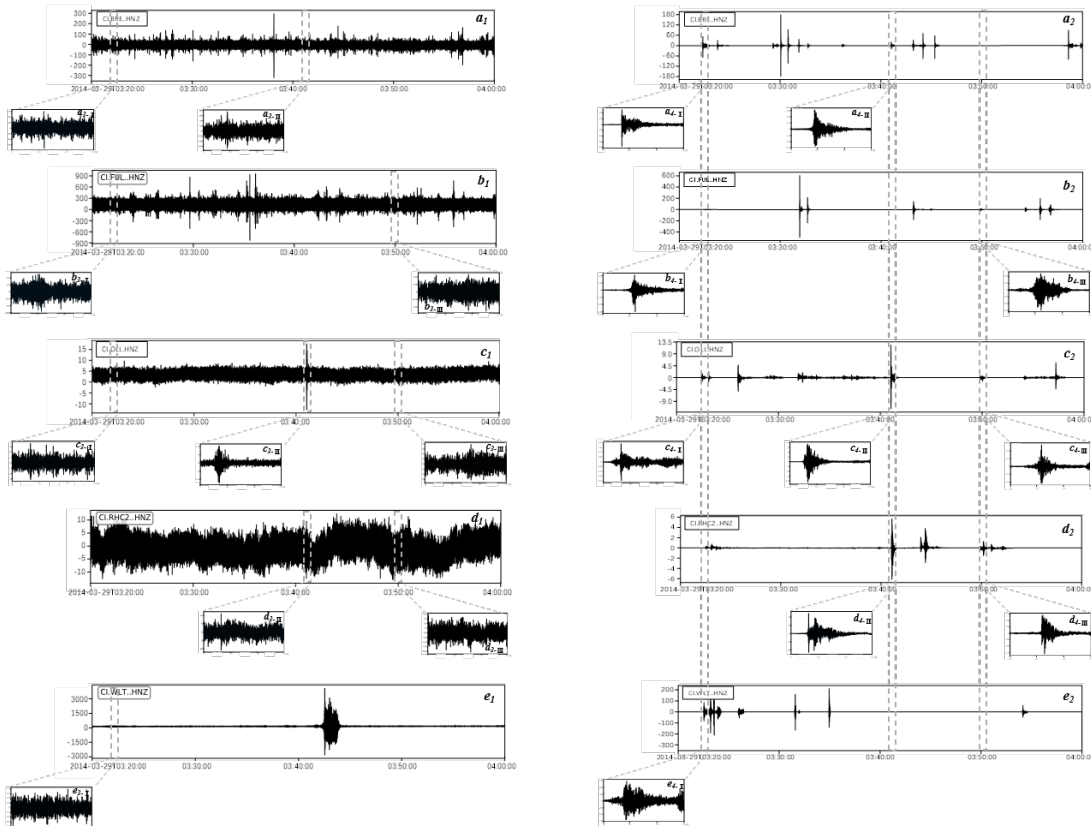

Figure 6. Application of UrbanDenoiser to the 40-minute seismograms (3:20 – 4:00 UTC, Mar. 29, 2014, vertical component) from the five stations of SCSN (Station CI.BRE, CI.FUL, CI.OLI, CI.RHC2 and CI.WLT). a1-e1 Raw seismograms. a2-e2 Denoised seismograms. a4-e4 Zoomed view of the denoised potential earthquake waveforms compared with the raw waveforms a3-e3.

Fig. 7 shows a comparison between the SNR of the denoised signals vs. non-denoised signals from Station CI.FUL for 102 events with -0.16 < M < 5.1. The SNR of the non-denoised data decreases rapidly with decreasing magnitude (black dots). UrbanDenoiser enhances the SNR for each event (red dots). Although the SNR of the denoised data decreases when the magnitude decreases, the SNR is consistently higher, and the trend is slower. On average, UrbanDenoiser enhances the SNR by about ten-fold, with the most dramatic improvement around M 1.5 - 3.8 (SNR: 2 - 100). This compares with a recently reported increase of a factor of 2-5 increase in SNR reported for denoising applied to more typical seismological settings[27].
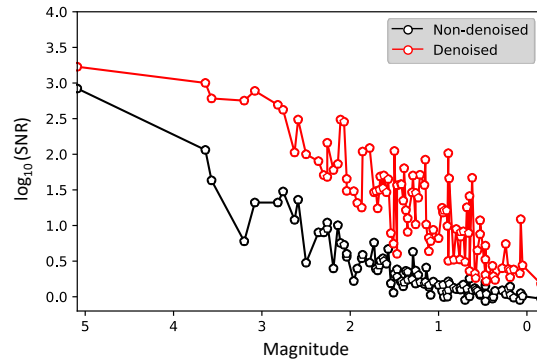

Figure 7. SNR of the denoised signals vs non-denoised signals from Station CI.FUL for 102 events varying between M -0.16 to M 5.1.

Discussion

The detection/location results contain the earthquake events only but exclude those large amplitude non-earthquake sources. Conventional detection methods detect sources of energy with amplitude that exceeds the detection threshold, but cannot differentiate earthquakes from other signals, such as waveforms generated by vibroseis or traffic. UrbanDenoiser can effectively suppress the high noise levels, though false positives and false negatives in denoised data should still be expected to occur and need to be assessed. The influence of false positives in denoised data can be effectively suppressed by using the dense array data for detection. False negatives occur when the seismic signal is too weak or when the target seismic phases and the training signal samples are not similar to the earthquake waveforms.

Although the aim of UrbanDenoiser is to separate earthquake signals from the urban noise, it could be extended to denoising for vibroseis waveforms by training it using high SNR signal samples from vibroseis events. In this case the vibroseis would be treated as the signals that we want to recover, while the earthquake signals would be grouped with the noise. Cleaner vibroseis signals should benefit the follow-on signal processing for seismic imaging.

For the most part, we do not have dense array deployments like Long Beach phase A and B available to generate a more complete earthquake catalog. The implementation of seismic monitoring relies on the isolated seismic instrument from the regional seismic network. The conventional Short Time Average/Long Time Average method can result in many false detections for phase identification such that it degrades the performance of phase association and event location. UrbanDenoiser can remove most of the noise bursts from the raw data and significantly increase the SNR for the seismic recordings in a single trace. This benefits the subsequent earthquake detection processing and should enhance the effectiveness of seismic monitoring by regional seismic networks in urban areas.

The earthquake detection results shown in Fig. 4 do not show evidence for widespread seismicity below 20 km in the upper mantle. We observe a weak tendency for events to follow the surface trace of the Newport-Inglewood Fault at 0 – 5 km. The seismicity at 5 – 15 km is more dispersed, which could be due to fault locking. The new seismic location result is consistent with our previous study[10], which found that at 15 – 20 km depth range, or roughly the root of the seismogenic zone, seismicity concentration is greater than at shallower depths. This could be due to the stress concentration near the seismic-aseismic transition. Based on our analysis, we conclude that earthquake detection and location following pre-processing using deep-learning to suppress noise should enable improved earthquake monitoring in urban environments.

Methods

Back-projection (BP) Imaging

We perform BP to image the earthquake location in two steps: (1) time-shift of each seismogram and (2) stacking. It can be expressed as[6]

$$stack_i(t) = \sum_{k=1}^{n} s_k\left(t - t_{ik}\right),\qquad(1)$$

where $s_k(t)$ is the seismogram recorded at the k-th station, t_{ik} is the calculated traveltime from the i-th grid point to the k-th station based on a known 3D velocity model, n is the number of stations, and $stack_i(t)$ is the stacked seismogram for the i-th grid-searching point.

We calculate the traveltime between each grid point and each geophone at the surface based on the Southern California Earthquake Center Community Velocity Model (CVM-H 11.9.1)[28] and store them in a traveltime lookup table for computational convenience. We perform a grid search over each potential source location within the imaging region. For each grid point, all the seismograms are time-shifted based on the corresponding traveltime, and the aligned seismograms are stacked to a single representative time series. The largest amplitude value along the time series is set as the amplitude value at that imaging point.

Seismic Data Processing and Back-projection Location with Continuous Data

We convert the original data in SEG-D format to NumPy format, decimate the time series from 500 Hz to 100 Hz, and process them with UrbanDenoiser. The denoised data are down sampled to 50 Hz, and band-pass filtered from 5-15 Hz. We normalize the data with their one-hour maximum value to suppress the influence of any strong spatially dependent residual noise level and calculate the envelope by smoothing the data with a three-point median window on the squared waveforms to reduce the sensitivity to the inaccuracy of the velocity model.

We perform BP for a 4.4 km (X) × 6 km (Y) × 25 km (Z) 3D imaging volume with a grid spacing of 200 m in each dimension. The geographic boundary of the imaging volume is shown as the red dashed rectangle in Fig. 1b. We perform BP as described above. We segment the shifted-and-stacked time series for each grid point into three-second time windows, and the maximum value within each time window is assigned as the BP value of this grid point. We thus obtain a 3D imaging volume for each three-second time window. If the maximum BP value through the whole space within a time window exceeds the detection threshold, we mark the corresponding grid point as a detection.

We set ten times the median absolute deviation (MAD) as the detection threshold for BP earthquake detection with the Long Beach dense array data. We fit the peak amplitude values from all time windows at each imaging point with a generalized extreme value (GEV) distribution. Under this detection threshold, we expect the number of false detections to be less than one per day.

References

1. Got, J.-L., Fréchet, J. & Klein, F. W. Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea. *J. Geophys. Res.* **99**, 15375 (1994).

2. Waldhauser, F. A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California. *Bulletin of the Seismological Society of America* **90**, 1353–1368 (2000).

3. Kao, H. & Shan, S.-J. The Source-Scanning Algorithm: mapping the distribution of seismic sources in time and space. *Geophysical Journal International* **157**, 589–594 (2004).

4. Davenport, K. K. *et al.* Aftershock imaging using a dense seismometer array (AIDA) after the 2011 Mineral, Virginia, earthquake. in *The 2011 Mineral, Virginia, Earthquake, and Its Significance for Seismic Hazards in Eastern North America* (Geological Society of America, 2015).

5. Ben-Zion, Y. *et al.* Basic data features and results from a spatially dense seismic array on the San Jacinto fault zone. *Geophysical Journal International* **202**, 370–380 (2015).

6. Beskardes, G. D. *et al.* A comparison of earthquake backprojection imaging methods for dense local arrays. *Geophysical Journal International* **212**, 1986–2002 (2018).

7. Inbal, A., Clayton, R. W. & Ampuero, J.-P. Imaging widespread seismicity at midlower crustal depths beneath Long Beach, CA, with a dense seismic array: Evidence for a depth-dependent earthquake size distribution: LB ARRAY SEISMIC MONITORING. *Geophys. Res. Lett.* **42**, 6314–6323 (2015).

8. Inbal, A., Ampuero, J. P. & Clayton, R. W. Localized seismic deformation in the upper mantle revealed by dense seismic arrays. *Science* **354**, 88–92 (2016).

9. Li, Z., Peng, Z., Hollis, D., Zhu, L. & McClellan, J. High-resolution seismic event detection using local similarity for Large-N arrays. *Sci Rep* **8**, 1646 (2018).

10. Yang, L., Liu, X. & Beroza, C. G. Revisiting Evidence for Widespread Seismicity in the Upper Mantle under Los Angeles. *Sci. Adv.* **7**, eabf2862 (2021).

11. Nazareth, J. J., & Hauksson, E. The seismogenic thickness of the Southern California crust, *Bull. Seismol. Soc. Am.* **94 (3)**, 940–960 (2004).

12. Bonar, D. & Sacchi, M. Denoising seismic data using the nonlocal means algorithm. *GEOPHYSICS* **77**, A5–A8 (2012).

13. Gaci, S. The Use of Wavelet-Based Denoising Techniques to Enhance the First-Arrival Picking on Seismic Traces. *IEEE Trans. Geosci. Remote Sensing* **52**, 4558–4563 (2014).

14. Mousavi, S. M., Langston, C. A. & Horton, S. P. Automatic microseismic denoising and onset detection using the synchrosqueezed continuous wavelet transform. *GEOPHYSICS* **81**, V341–V355 (2016).

15. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).

16. DeVries, P. M. R., Viégas, F., Wattenberg, M. & Meade, B. J. Deep learning of aftershock patterns following large earthquakes. *Nature* **560**, 632–634 (2018).

17. Perol, T., Gharbi, M. & Denolle, M. Convolutional neural network for earthquake detection and location. *Sci. Adv.* **4**, e1700578 (2018).

18. Ross, Z. E., Yue, Y., Meier, M., Hauksson, E. & Heaton, T. H. PhaseLink: A Deep Learning Approach to Seismic Phase Association. *J. Geophys. Res. Solid Earth* **124**, 856–869 (2019).

19. Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y. & Beroza, G. C. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat Commun* **11**, 3952 (2020).

20. Zhu, W., Mousavi, S. M. & Beroza, G. C. Seismic Signal Denoising and Decomposition Using Deep Neural Networks. *IEEE Trans. Geosci. Remote Sensing* **57**, 9476–9488 (2019).

21. Snover, D., Johnson, C. W., Bianco, M. J. & Gerstoft, P. Deep Clustering to Identify Sources of Urban Seismic Noise in Long Beach, California. *Seismological Research Letters* **92**, 1011–1022 (2021).

22. Zhu, W., Mousavi, S. M. & Beroza, G. C. Seismic signal augmentation to improve generalization of deep neural networks. *Advances in Geophysics* **61**, 151–177 (2020).

23. Meng, H., Ben-Zion, Y. & Johnson, C. W. Detection of random noise and anatomy of continuous seismic waveforms in dense array data near Anza California. *Geophysical Journal International* **219**, 1463–1473 (2019).

24. Ross, Z. E., Trugman, D. T., Hauksson, E. & Shearer, P. M. Searching for hidden earthquakes in Southern California. *Science* **364**, 767–771 (2019).

25. Zhu, W. & Beroza, G. C. PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method. *Geophysical Journal International* **216**, 261–273 (2018).

26. Shaw, J. H. *et al.* Unified Structural Representation of the southern California crust and upper mantle. *Earth and Planetary Science Letters* **415**, 1–15 (2015).

27. Tibi, R., Hammond, P., Brogan, R., Young, C. J. & Koper, K. Deep Learning Denoising Applied to Regional Distance Seismic Data in Utah. *Bulletin of the Seismological Society of America* **111**, 775–790 (2021).

28. Plesch, A., C. Tape, J. R. Graves, P. Small, G. Ely, and J. H. Shaw (2011), Updates for the CVM-H including new representations of the offshore Santa Maria and San Bernardino basin and a new Moho surface, paper presented at Southern California Earthquake Center Annual Meeting, Palm Springs, Calif.